基于最优近似粗糙集的属性约简 *

罗来鹏，刘二根，范自柱

(华东交通大学 理学院，南昌 330013)

摘　要：为了更好地获取由边界域产生的不确定性规则知识，提出最优近似粗糙集的属性约简方法，为此首先给出了近似空间上粗糙集最优近似集的判定与计算，然后引入最优近似分布协调集、最优近似分布约简概念。讨论了 Pawlak 属性约简、分布约简、最优近似分布约简之间关系，最后得到在协调决策表中它们是等价的，在不协调决策表中最优近似分布约简是分布约简子集。从 UCI 数据集选取 5 个数据集进行实验，结果表明基于最优近似分布约简能够获得较少的约简属性。

关键词：粗糙集；属性约简；相似度；最优近似集；分布约简；最优近似分布约简

中图分类号：TP391　　doi: 10.3969/j.issn.1001-3695.2017.12.0856

Attributes reduction based on optimal approximation set of rough set

Luo Laipeng, Liu Ergen, Fan Zizhu

((*School of Sciences*, *East China Jiaotong University*, *Nanchang* 330013, *China*)

Abstract: To obtain the knowledge of uncertainty rules generated by the boundary region, the paper developed an attribute reduction algorithm based on the optimal approximate rough set . Firstly this paper presented the judgement and calculation of the optimal approximation set of rough set in approximate space ; secondly, it introduced the concepts of the optimal approximate distribution coordination sets and the optimal approximate distribution reduction, and discussed. the relations among Pawlak attributes reduction, distribution reduction and distribution reduction based on the optipal approximation set of rough set ; finally, it obtained the conclusion that they are equivalent in the coordination decision table and the distribution reduction set based on the optipal approximation set of rough set is subset of the distribution reduction sets. Experiment results on datasets from UCI show that the attribute reduction algorithm based on the optimal approximate rough set can achive less set of attributes.

Key words: rough set; attributes reduction; similarity degree; optimal approximation set ; distribution reduction; optipal approximation distribution reduction

0　引言

属性约简是粗糙集理论[1]的应用核心,基于它的知识获取往往存在不确定性。一方面是因为近似空间上知识的粒太大而造成概念的不确定性；另一方面是由于近似空间上粗糙集边界造成的系统不确定性[2]。对于一个给定的近似空间，概念的不确定性往往无法改变，因此如何更好挖掘系统不确定性所蕴涵的知识一直是众多研究者关注的焦点。当前比较典型的研究模型有变精度粗糙集模型和概率粗糙集模型.变精度粗糙集模型主要是通过引进粗糙性测度或近似精度来实现。米据生等人讨论了变精度粗糙集模型[3~5],并利用它进行属性约简,取得了较好的效果[10]，并由此发展了粗糙集的三支决策[6~8]。Yao 和 Ziarko 等人[9,10]结合概率论和包含度提出了概率粗糙模型,也取得了较好的理论成果。这两个模型都是在基本粗糙集模型的基础上,通过引入一个在区间（0.5,1]上阈值作为分类正确率而建立。阈值的具体取值对模型的属性约简结果具有直接影响,在实际应用中很难确定一个最优的阈值.此外这两个模型目标概念的表示仍然使用上、下近似的表示方法.那么在目标概念的上、下近似集之间哪个集合更为近似目标概念呢？近几年张清华教授对这个问题展开了研究，提出目标概念新的刻画，即粗糙集的最优近似[11]。文献[12,13]以集合的相似度大小为准则,通过模糊截集的方法,在目标概念的上、下近似集之间构造一个隶属度大于或等于 λ 集合 $R_\lambda(X)$,讨论 $R_\lambda(X)$ 和目标概念两个边界集作为目标概念最优近似集的条件,并就最优近似集的问题作了进一步的研究得到一些新的理论成果.但是对于一个具体的目标概念集未讨论如何求解最优近似集，尤其基于粗糙集最优近似集表示下属性约简问题.针对上述问题本论文提出粗糙集的最优近似集新判定与计算,并在此基础上提出最优近似粗糙集的属性约

收稿日期：2017-12-30；**修回日期：**2018-02-27　**基金项目：**国家自然科学基金资助项目（61472138）

作者简介：罗来鹏（1973-），男，江西吉水人，副教授，硕士，主要研究方向为粗糙集与粒计算（luolp789@163.com）；刘二根（1965-），男，教授，硕士，主要研究方向为优化与计算；范自柱（1975-），男，教授，博士，主要研究方向为智能计算与图像处理.

简算法,并讨论该算法约简结果与其他相关算法约简结果之间关系。

1　相关概念

定义[1]1　设 U 是有限论域,R 是 U 上的等价关系,由此构成一个近似空间为(U,R),等价关系 R 决定的等价类也称为近似空间 (U,R) 上基本知识，表示为 U/R= $\{[x_i]_R | x_i \in U\}$ 其中 $[x_i]_R = \{x_j | (x_i, x_j) \in R\}$。

定义[1]**2**　(U,R)为近似空间，$X \subseteq U$ 为目标概念，则 $\underline{R}(X) = \{[x]_R | [x]_R \subseteq X\}$，$\overline{R}(X) = \{[x] | [x]_R \cap X \neq \phi\}$ 分别称为目标概念 X 关于近似空间(U,R)下近似集、上近似集。$BN_R(X) = \overline{R}(X) - \underline{R}(X)$，$NEG_R(X) = U - \overline{R}(X)$ 分别称为目标 X 在近似空间(U,R)上边界集和负域。

定义 3[1] (U,R)为近似空间,目标概念 $X \subseteq U$,若 $\underline{R}(X) = \overline{R}(X)$，则称目标概念 X 在近似空间(U,R)上是可以定义的精确集,否则为粗糙集。

如果 X 是近似空间(U,R)上的粗糙集,一般用它的上、下近似集作为边界来表示即 $(\underline{R}(X), \overline{R}(X))$。

定义[14]4　设(U,R)为近似空间，$X \subseteq U, Y \subseteq U$，若映射 S: $U \times U \to [0,1]$,即 S(X,Y)满足以下条件:

a)有界性。$0 \le S(X,Y) \le 1$,特别地 S(X,X)=1，S(X,Y)=0 当且仅当 $X \cap Y = \phi$。

b)对称性。$S(X,Y) = S(Y,X)$。

c) $X \subseteq Y \subseteq Z$, $S(X,Y) \ge$ S(X,Z),S(Y,Z) ≥S(X,Z)。

称 S 为集合 X,Y 的相似度.

对于经典集合，两个集合的相似定义度可定义为 $S(X,Y) = \frac{|X \cap Y|}{|X \cup Y|}$,其中 $|\bullet|$ 表示集合基数,下文都是这种约定.

2　基于最优近似集的属性约简

2.1　粗糙集的最优近似集

设 X 是近似空间(U,R)的目标概念，令 $\alpha = \frac{|\underline{R}(X)|}{|X|}$，$\beta = \frac{|X|}{|\overline{R}(X)|}$

$B_X = \{[x]_R | 0 < \frac{|[x]_R \cap X|}{|[x]_R|} < 1, x \in U\}$，

$C_X = \{[x]_R | \frac{|[x]_R \cap X|}{|[x]_R - X|} \ge \alpha, [x]_R \in B_X\}$，则 B_X 是 X 在近似空间（U,R）上的边界域,α, β 分别为 X 与它的下近似集、上近似集的相似度,且 $C_X \subseteq B_X$。对于任意 $A \subseteq B_X$，$\underline{R}(X) \cup A$ 都是 X 的近似集,因此 X 在近似空间上的近似集合会有很多个。

定义 5　设 X 在近似空间（U,R）上所有近似集为 X^{apps}，若存在 $X_1 \in X^{apps}$，使得对于任意的 $X_2 \in X^{apps}$，有 S(X,X1)≥S(X,X2)则称 X1 为 X 在（U,R）上最优近似集。

接下来讨论近似空间上一个目标概念的最优近似计算

定理 1 X 为近似空间(U,R)上的目标概念,若 $\Phi \subset C_X \subset B_X$,则 X 最优近似集为 $\underline{R}(X) \cup C_X$。

证明　对于任意 $[y]_R \in C_X$ 有：$\frac{|[y]_R \cap X|}{|[y]_R - X|} \ge \alpha$，不妨设 $C_X = \bigcup_{i=1}^{k} [y_i]_R$，且 $\frac{|[y_1]_R \cap X|}{|[y_1]_R - X|} \le \cdots \le \frac{|[y_k]_R \cap X|}{|[y]_k - X|}$，根据不等式性质有：

$$S(\underline{R}(X), X) \le \frac{|\underline{R}(X)| + |[y_1]_R \cap X|}{|X| + |[y_1]_R - X|} \le \cdots \le \frac{|\underline{R}(X)| + \sum_{i=1}^{k} |[y_i]_R \cap X|}{|X| + \sum_{i=1}^{k} |[y_i]_R - X|}$$

而对于任意 j(j<k)有：

$$S(\underline{R}(X) \cup [y_1]_R \cup \cdots \cup [y_j]_R), X) = \frac{|(\underline{R}(X) \cup [y_1]_R \cup \cdots \cup [y_j]_R) \cap X|]}{[(\underline{R}(X) \cup [y_1]_R \cup \cdots \cup [y_j]_R) \cup X[}$$

$$= \frac{|(\underline{R}(X) \cap X) \cup ([y_1]_R \cap X) \cup \cdots \cup ([y_j]_R \cap X)|}{|\underline{R}(X) \cup X \cup [y_1]_R \cup \cdots \cup [y_j]_R|} = \frac{|\underline{R}(X)| + \sum_{i=1}^{j} |[y_i]_R \cap X|}{|X \cup [y_1]_R \cup \cdots \cup [y_j]_R|}$$

$$= \frac{|\underline{R}(X)| + \sum_{i=1}^{j} |[y_i]_R \cap X|}{|X| + \sum_{i=1}^{j} |[y_i]_R - X|}$$

即 $S(\underline{R}(X), X) \le S((\underline{R}(X) \cup [y_1]_R), X) \le \cdots \le S((\underline{R}(X) \cup B_X), X)$,

所以 X 最优近似集为 $\underline{R}(X) \cup C_X$。

根据上述结论得到以下两个结论：

性质 1　若 Cx=ϕ ,则 X 最优近似集为下近似集 $\underline{R}(X)$。

性质 2　若 $C_X = B_X$,则 X 最优近似集为上近似集 $\overline{R}(X)$。

2.2　基于最优近似的分布约简

定义 6　决策系统 DS=(U,C∪D,V,f),其中 U 为有限论域,C 为条件属性集,D 为决策属性集,V 为属性值集合，f 为信息函数,且 U/C={C1,C2,…Cm},U/D={D1,D2,Dn}。

接下来讨论它的最优近似约简。

记 Di 在 U/C 上的最优近似集表示为: $D_C^{(D_i)}$,Ci 所在的最优近似集表示为 $D_C(C_i)$ 并称 $D_C^* = \{D_c^{(D_1)}, D_C^{(D_2)}, \cdots, D_C^{(D_n)}\}$ $(D_i \in U/D)$ 为 D 在 C 上的最优近似分布。

定理 2　设决策系统 DS=(U,C∪D,V,f)是协调的,那么对于任意 i 有 $D_C^{(D_i)} = \underline{C}(D_i)$,反之不成立。

证明　若决策系统 DS 是协调的,对于任意 $D_i \in U/D$ 有 $B_{D_i} = \Phi$，所以有 $D_C^{(D_i)} = \underline{C}(D_i)$，反之不成立，如 U/C={{x1,x2},{x3,x4,x5},{x6,x9},{x7,x8}},U/D={{x1,x2,x3},{x4,x6,x9},{x5,x7,x8}},对于任意 Di∈U/D 在 U/C 上的最优近似集为其下近似集,显然这个划分对应的的决策系统非协调。

定义 7　决策系统 DS=(U,C∪D,V,f),其中 C 为条件属性集,D 为决策属性集,$B \subseteq C$,若 $D_C^* = D_B^*$,则称 B 为最优近似分布协调集.若 B 是最优近似分布协调集且 B 的任何真子集都不是最优近似分布协调集,则称 B 为最优近似分布约简。

定义 8　决策系统 DS=(U,C∪D,V,f),其中 C,D 为条件属性集，决策属性集,C,D 决定的等价类为：U/C={C1,C2,…Cm},U/D={D1,D2, …Dn},用 $f_k(C_i)$ 表示属性 a_k

关于 C_i 中对象的取值,定义

$$D(C_i,C_j)=\begin{cases}\{a_k\in C: f_k(C_i)\neq f_k(C_j) & D_C(C_i)\neq D_C(C_j)\} \\ C & D_C(C_i)=D_C(C_j)\end{cases}$$

则 $D(C_i,C_j)$ 为 C_i 和 C_j 可辨识属性集。

定义 9 决策系统 DS=(U,C∪D,V,f),其中 C,D 分别为条件属性集,决策属性集,U/C={C1,C2,…Cm},U/D={D1,D2,…,Dn}, $D=\{D(C_i,C_j),i,j\le m\}$ 为最优近似分布可辨识矩阵。

根据辨识矩阵属性约简方法,得到如下定理:

定理 3 设目标信息系统 DS=(U,C∪D,V,f),辨识公式 $\wedge\{\vee(a_k\in D(C_i,C_j):i,j\le m\}$ 的极小析取范式为决策目标系统的最优近似分布约简。

具体证明可参考相关文献[15]。

2.3 基于最优近似集的属性约简算法描述

决策系统 DS=(U,C∪D,V,f),其中 C,D 分别为条件属性集,决策属性集且 U/C={X1,X2,…Xm},U/D={D1,D2,…， Dn},最优近似分布约简步骤如下:

a)计算 $D_C^{(D_i)}(i=1,2,\cdots,n)$,得到 D_C^*,具体为:

(a)对于任意 $D_i\in U/D$,计算 $\underline{R}(D_i)$, $\alpha=\frac{\underline{R}(D_i)}{|D_i|}$ 。

(b)计算

$$B_{D_i}=\{X_k\,|\,0<\frac{|X_k\cap D_i|}{|X_k|}<1, X_k\in U/C\}, C_{D_i}=\{X_k\,|\,\frac{|X_k\cap D_i|}{|X_k-D_i|}\ge\alpha, X_k\in B_{D_i}\}$$

（c）判断与求解。若 $C_{D_i}=\Phi$, D_i 最优近似集为下近似集 $\underline{R}(D_i)$ ；若 $C_{D_i}=B_{D_i}$,则 D_i 最优近似集为上近似集 $\overline{R}(D_i)$ ；$\Phi\subset C_{D_i}\subset B_{D_i}$,则 D_i 最优近似集为 $\underline{R}(D_i)\cup C_{D_i}$ 。

b)计算不同类之间基于最优近似集的辨识属性集 $D(C_i,C_j)(i,j\le m)$ 并构造最优近似分布分辨矩阵。

c)求出最优近似分布矩阵极小析取范式,得到最优近似分布。

3 最优近似分布约简与其他约简之间的关系

定义 10[1] 决策系统 DS=(U,C∪D,V,f),C 为条件属性集,D 为决策属性集,若 $B\subseteq C$, $pos_B(D)=pos_C((D)$,且 B 是 C 的 D 独立子集,那么 B 为 Pawlak 属性约简也称正域约简.

记 $D(D_j/[x]_C)=\frac{|D_j\cap[x]_C|}{|[x]_C|}(j\le n)$ ，则 D 为 D_j与$[x]_C$ 的包含度,

$\mu_C(x)=(D(D_1/[x]_C),D(D_2/[x]_C),\cdots,D(D_n/[x]_C),(x\in U)$ ，则称 $\mu_C(x)$ 为关于属性 C 在决策系统 DS 中的广义分布函数[2]。

定义 11[2] 决策系统 DS=(U,C∪D,V,f),其中 C,D 分别为条件属性集,决策属性集,C,D 决定的等价类为 U/C={C1,C2,⋯Cm},U/D={D1,D2,Dn}, $B\subseteq C$,若对于任意 $x\in U$ 有 $\mu_C(x)=\mu_B(x)$,则称 B 是分布协调集。若 B 是分布协调集且 B 的任何真子集不是分布协调集,则称 B 是分布约简。

定理 4 决策系统 DS=(U,C∪D,V,f),其中 C 为条件属性集,D 为决策属性集,若决策表是协调的,那么决策系统的 Pawlak 属性约简、分布约简与最优近似分布约简等价。

证明 若 B 是协调决策表 DS 的 Pawlak 约简,则对于任意 $D_i\in U/D$ ，有 $\underline{C}(D_i)=\underline{B}(D_i)=D_i$ ，即有 $pos_C(D)=pos_B(D)$ ，和 $D_C^*=D_B^*$ 且对于任意 x∈U, $\frac{|D_i\cap[x]_C|}{|[x]_C|}=\frac{|D_i\cap[x]_B|}{|[x]_B|}$ 即 $\mu_C(x)=\mu_B(x)$ 故 B 也是分布约简和最优近似分布约简。

而当决策系统是非协调时,最优近似分布约简与分布约简不等价。

定理 5 设决策系统 DS=(U,C∪D,V,f),其中 C 为条件属性集,D 为决策属性集, $B\subseteq C$,若 B 是决策系统的分布约简，则 B 是系统的最优近似分布协调集。

证明 对于任意 $D_i\in U/D$, $\mu_C(x)=\mu_B(x)$ ，（x∈U),即 $\frac{|D_i\cap[x]_B|}{|[x]_B|}=\frac{|D_i\cap[x]_C|}{|[x]_C|}$,由等比性质得到 $\frac{|D_i\cap[x]_B|}{|[x]_B|-|D_i\cap[x]_B|}=\frac{|D_i\cap[x]_C|}{|[x]_A|-|D_i\cap[x]_C|}$,由集合性质得到 $\frac{|D_i\cap[x]_B|}{|[x]_B-D_i|}=\frac{|D_i\cap[x]_A|}{|[x]_A-D_i|}$,所以有 $D_C^*=D_B^*$,故 B 是最优近似协调集.

分布约简是最优近似协调集,但是未必是最优近似分布约简,最优近似分布约简是分布约简子集。比如决策系统 DS=(U,C∪D,V,f),其中

U={x1,x2,x3,x4,x5,x6},C={a1,a2,a3,a4},D={d},U/C={C1,C2,C3,C4},U/d={D1,D2}D1={x1,x5,x6},D2={x2,x3,x4}C1={x1},C2={x2},C3={x3,x5,x6},C4={x4}

它的分布约简与最优近似约简如下:

μ_C (x1)=(1,0), μ_C (x2)=(0,1), μ_C (x3)= μ_C (x5)= μ_C (x6)=(2/3,1/3), μ_C (x4)=(0,1)

D(C1,C2)={a1,a2,a3,a4},D(C1,C3)={a1,a2},D(C1,C4)={a1,a2,a3},D(C2,C3)={a3,a4},D(C3,C4)={a3}

辨识公式为

(a1 ∨ a2 ∨ a3 ∨ a4) ∧ (a1 ∨ a2) ∧ (a1 ∨ a2 ∨ a3) ∧ (a3 ∨ a4)a3=(a1a3)∧(a2a3),即 a1a3,a2a3 为目标系统的两个分布约简.

下面给出最优近似分布约简:

$D_C^{(D_1)}$ =(C1,C3},

$D_C^{(D_2)}$ =(C2,C4}D(C1,C2)={a1,a2,a3,a4},

D(C1,C4)={a1,a2,a3},D(C2,C3)={a3,a4},D(C3,C4)={a3}

辨识公式为

(a1∨a2∨a3∨a4)∧(a1∨a2∨a3)∧(a3∨a4)a3=a3 即是说 a3 为目标系统的唯一的最优近似分布约简.

综合得到:分布约简是最优近似协调集，但不是最优近似分布约简,最优近似分布约简是分布约简的子集.

根据定理 1,还可以得到如下结论:设决策系统 DS=(U,C∪D,V,f),其中 C 为条件属性集,D 为决策属性集，对于任意的 $D_i \in U/D$若 $D_C^{(D_i)} = \underline{C}(D_i)$,那么最优近似分布约简变为为正域约简或下近似约简,对于任意的$D_i \in U/D$，若 $D_C^{(D_i)} = \overline{C}(D_i)$,最优近似约简即变为上近似约简。这不仅进一步验证了上述结论，同时可以发现最优近似约简实际上是一种根据系统知识之间的关系而选择不同属性约简的综合，能更好地刻画了粗糙集不需要先验知识的特性。

4 实验

为了更好地说明本文所得到的结论，选用了 UCI 数据库中的 5 个数据集: Mushroom,Vote,Vehicle,Gene，Ticdata2000，就分布约简，最优近似约简的约简结果进行对比分析，用$|C_m|$表示约简后属性个数,采用 MATLAB 编程计算得到如表 1 所示结果。

从实验结果可以看出,在第一个数据集分布约简与最优近似分布约简的约简结果是一样，那是因为此时的最优近似约简和分布约简就是下近似约简，这也验证了最优近似约简与分布约简之间存在的关系，后四个数据集最优近似分布约简比分布约简得到更少的属性约简集,这些完全符合本文得到的理论结果。从压缩的比例来看，两个算法对这 5 个数据集对属性压缩比例如表 2 所示。

表 1 不同数据集的约简结果比较

Datasets	$\|U\|$	$\|C\|$	分布约简($\|C_m\|$)	最优近似分布约简($\|C_m\|$)
Mushroom	8124	22	5	5
Vote	435	16	10	8
Vehicle	946	18	11	9
Gene	3190	60	11	9
Ticdata2000	5822	85	24	20

表 2 不同数据集的属性压缩比例比较

Datasets	$\|C\|$	分布约简压缩比例	最优近似分布约简压缩比例	压缩比例提高百分比
Mushroom	22	77.3%	77.3%	0%
Vote	16	37.5%	50%	12.5%
Vehicle	18	39%	50%	11%
Gene	60	81.7%	85%	3.3%
Ticdata200	85	71.8%	76.5%	4.7%

从表 2 可以发现最优近似分布约简对属性压缩比例要高，后面两个数据集之所以提高百分比不大，那是因为属性基数太大，属性个数改变对百分比的影响要小，但是从约简绝对个数还是很显著。

5 结束语

面对不确定性系统的属性约简,目前比较流行的方法是变精度粗糙集和概率粗糙集模型,但是这两种方法都存在着参数如何设置问题.文提出的基于最优近似属性约简方法虽然也属于分布约简范畴,但是弱化了参数的设置。并且从约简结果之间关系可以看出基于最优近似集的属性约简不仅具有更少的属性约简个数，而且具有一定的泛化性能,为研究决策系统之间属性关系拓展了思路.那么粗糙集的最优近似、变精度近似与概率近似之间存在怎样关系，笔者认为这是一个值得研究问题，这对于进一步探索不同属性约简结果之间关系具有重要作用。

参考文献：

[1] Pawlak Z. Rough sets [J]. Computer and Information Science, 1982, 11 (5): 341-356.

[2] 张文修, 梁怡, 吴伟志. 信息系统与知识发现 [M]. 北京: 科学出版社, 2003.

[3] Mi Jusheng, Wu Weizhi, Zhang Wenxiu. Approaches to knowledge reduction based on variable precision rough set model [J]. Information Sciences, 2004, 159 (3-4): 255−272.

[4] 张贤勇. 精度与程度逻辑差双量化粗糙集模型的属性约简 [J]. 系统工程理论与实践, 2015, 35 (11): 2925-2931.

[5] 曾子林, 张宏军, 张睿, 等. 变精度粗糙集的逻辑解释及其约简 [J] 计算机应用研究, 2013, 30 (5): 1385-1387.

[6] 张鑫, 李续武, 路艳丽, 等. 基于不完备区间值信息系统的决策粗糙集 [J]. 计算机应用研究, 2017, 34 (1): 110-112.

[7] Feng Qinrong, ZhouYing. Soft discernibility matrix and Its applications in decision making [J]. Applied Soft Computing, 2014, 24 (11): 749-756.

[8] 李美争, 王国胤. 三支近似概念格中基于对象-概念辨识矩阵的属性约简方法 [J]. 控制与决策, 2016, 31 (10): 779-784.

[9] Yao Yiyu. Probabilistic rough set approximations [J]. Approximate Reasoning, 2008, 49 (2): 255-271.

[10] Ziarko W. Probabilistic approach to rough sets [J]. Approximate Reasoning, 2008, 49 (2): 272-284.

[11] 张清华, 王国胤, 肖雨. 粗糙集的近似集 [J]. 软件学报, 2012, 23 (7): 1745−1759.

[12] 张清华, 薛玉斌, 胡 峰, 等. 粗糙集近似集不确定性研究 [J]. 电子学报, 2016, 44 (7): 1574-1580.

[13] 张清华, 薛玉斌, 王国胤. 粗糙集的最优近似集 [J]. 软件学报, 2016, 27 (2): 295-308.

[14] 袁修久, 张文修. 模糊粗糙集的包含度和相似度 [J]. 模糊系统与数学, 2005, 19 (1): 111-115.